AD________________

Award Number: DAMD17-99-1-9374

TITLE: The CAD Method for Microcalcification Detection: Independent of Sensor and Resolution

PRINCIPAL INVESTIGATOR: Wei Qian, Ph.D.

CONTRACTING ORGANIZATION: University of South Florida
Tampa, Florida 33620-7900

REPORT DATE: July 2000

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2000	Annual Summary (1 Jul 99 - 30 Jun 00)

4. TITLE AND SUBTITLE
The CAD Method for Microcalcification Detection: Independent of Sensor and Resolution

5. FUNDING NUMBERS
DAMD17-99-1-9374

6. AUTHOR(S)
Wei Qian, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of South Florida
Tampa , Florida 33620-7900

E-MAIL:
wquian@hsc.usf.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution unlimited	

13. ABSTRACT *(Maximum 200 Words)*

The complishments of the first year's task in this project are summarized as follows:

(1). Data base collection and truth file establishment for three sensors
(2). Breast area segmentation for the mammograms
(3). Basic algorithm design including nonlinear bank filter design, implementation and segmentation algorithm design and testing.

These achievements provide an very important background for the future implementation of our long term aims of the project, which is the development of a more generalizable, automatic, and robust computer assisted diagnostic (CAD) method, for microcalcification cluster (MCC) detection using digital mammography; suitable for both digitized film and direct X-ray sensors. The *specific aim* of this proposal is to evaluate the feasibility of developing an *entirely new class of adaptive multiresolution CAD method,* implemented on *non-linear* filter banks, and compare this method to a wavelet transform method, implemented on *linear* filter banks, as previously reported.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	11
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

I. Introduction:

This project proposed a very novel approach for CAD for MCC's namely, an *adaptive CAD* method that allows for *variations in the image from different sensors and is independent of resolution.* Preliminary data are presented that show this dependence and how the CAD method can be modified to find a solution for a generalized CAD method as required for multi center clinical trials. *The performance of the CAD method in its present form is better than nay reported to date allowing for data base dependence.* An entirely new class of CAD method is proposed, fundamentally different in design compared to current CAD methods and the result of several years of research where the mathematical algorithms are designed for digital mammography. This proposal is not simply a recommendation of using a specific algorithm for this purpose, but is a very systematic approach where each CAD module is designed from the principles of signal processing. This report includes 3 sections, **(1). Data base collection and truth file establishment for different sensors, (2). Preprocessing for breast area segmentation, and (3) Basic algorithm design.** These works provide an very important background for the future implementation of our long term aims of the project, which is the development of a more generalizable, automatic, and robust computer assisted diagnostic (CAD) method, for microcalcification cluster (MCC) detection using digital mammography; suitable for both digitized film and direct X-ray sensors. Some of our progress is sumarized in our recent publications [1, 2, 3, 4].

BODY OF REPORT

I. Data base collection and truth file establishment for three sensors

In order to make the CAD algorithms for clinical use, the extensive tests of our algorithms are necessary for larger database *from different digitizers with different resolutions. The films of* the two sets of 200 views was digitized with two different digitizers: (a) a CCD-based system (ImageClear R3000, DBA Systems, Inc, Melbourne, FL) *maximum resolution at* 30 *m and* 16 bits and (b) a LUMISCAN 85 (Lumisys, Sunnyvale, CA) at maximum resolution, at 50 m, and 12 bits. Digitized images will be acquired at 50 microns with one set transformed to 100microns respectively. This is a resolution change only, not ageometric and intensity change that modifies the gray scale histogram where we havealready tested and found adequate for MCC detection at different resolutions. Our customdesigned workstation can accommodate both resolutions. Hence a total of 1200 images will be generated from 3 different digital sources.
A direct digital mammography system from General Electric was installed at the H. Lee Moffitt Cancer Center and Research Institute (MCCRI) at USF in June 1998. We have access to this system through the NCI (P30 Funded center) program structure at Moffitt where the imaging program collaborates with the mammography diagnostic and screening program. The system will be evaluated and used clinically in the diagnostic breast cancer program of MCCRI. Clinical protocols (two views each breast) will require that direct digital images will have a spatial resolution of 100 m and 12 bits per pixel, and will be evaluated in a clinical trial to show its equivalency to a standard screen/film system (Contour mammography system, Bennett, Copiague, NY). It is expected that at least 1,000 women per year will participate in the clinical trial and undergo both screen/film and direct digital mammography. Hence, about 1000 cases will have both film and digital mammograms by June 2000. These cases will be our pool from which two sets of 200 images will be selected for this study, the first as a training set, the second as a testing set. Thses sets will include the following: (a) 50 views with no findings (benign or malignant) which will remain as such for at least two years, i.e. their negative nature will be confirmed at the end of year 2000, with a follow-up mammogram. These will constitute our "normal cases", (b) 150 views with masses (benign and malignant). These cases will be selected consecutively and constitute our "abnormal cases". Based on current statistics from the breast cancer program of MCC, it is expected that masses of various shapes (round, oval, lobular, irregular), various margins (circumscribed, microlobulated, obscured, indistinct, spiculated), and various densities (high, equal, low, fat containing) will be included . The contents of each set will be evaluated once the desired number of images is accrued. *Accrual will be extended if all* major mass types are not adequately represented in the first 150 mass cases. Normal and abnormal cases will have radiology reports *following BIRADS; reports for the digital and* screen/film mammograms of the same patient will be similar. Abnormal cases will have electronic truth files with the location, size, and margins of *the mass for all views indicated by* an expert, specifically for both sensors to accommodate anticipated changes due tobreast positioning and compression. *Electronic truth files are important in the evaluation process to* allow a comparison of segmentation methods. The pathology of all the abnormal cases will be confirmed *during the same time period as the accrual.*

II. Preprocessing for breast area segmentation

1. The basic idea of design and improvement of our preprocessor

Except a few mammogram digital images of high quality, most of them suffer from more or less extrinsic signals such as images of edges of original X-ray pictures, or large bright regions caused by the cutting of them, various notes made by doctors, uneven illumination, blurred edges due to the exposure of pictures. Sometimes such extrinsic signals in the digital mammogram images may affect the detection result seriously.

We developed an algorithm to erase the extrinsic signals. They were helpful for the detection but still not satisfied. So we tried to find the more efficient way to do this work.

The mammogram images has their own characteristics. Some of them are below. They usually consist of a connected mammogram region with brighter inside area and darker boundary area. Except going extremes, most extrinsic signals are located outside the mammogram region, therefore removing such kind of extrinsic signals is nothing but letting our CAD program ignore those signals outside the mammogram region. Also, we have two approaches to do this job. One is to find the difference between the intrinsic signals and extrinsic signals. Another is to separate the outside region from the mammogram region. In short word, the first method is like taking medicine, the second is like taking operation. Many of preprocessor designed to remove extrinsic signals mainly by filtering images belongs to the first method. the advantage of filtering is that a lot of filters have been extensively studied and are ready to be applied stably, the extrinsic signals in the mammogram can be erased or suppressed. The disadvantage is filtering on both useful intrinsic signals and useless extrinsic signals, some useful intrinsic signal may be lost, some strong extrinsic signals can't be cleaned and still affect the detection result. The advantage of the second method is that CAD could gets rid of signals outside the mammogram region completely and keep the all information in the mammogram, the disadvantage is the extrinsic signals in the mammogram is remained and we need thresholds for stable separating regions. Our First preprocessor belongs to the first method. Our Second preprocessor(this one) belongs to the second method. To find a method to separate the mammogram region from the given digital image without filtering, is equivalent to find the boundary of mammogram region. The gray values of boundary of the mammogram regions depends on the different database of the images. For example, the gray values have a big jump around boundaries of images of Lumisys, while the jump is almost not visible around boundaries of images of DBA. we propose a novel method to find boundaries. The main idea of our method is not to chain existing edges to search the wanted boundaries, but to develop a new single contour, which has no need of chaining and grouping, in other word, the contour works much as a knife to cut the wanted mammogram region out of the given image.

2. The construction steps of our preprocessor

Several steps are needed to form our contour.

2.1 Thresholding globally

This step consists of minimal and maximal gray values of the given image evaluating and adjusting, threshold estimating for global developing of a contour. We choose two thresholds: lower- threshold and upper-threshold for evaluation which ignores all pixels with gray value below lower-threshold and above upper-threshold. We may choose different approaches to get these thresholds, and we choose jumping of mean values as gray value changes to search them. The reason of doing this way is this jumping is a statistic amount, not sensitive to images unless going extremes. The test results shows that this method is easy to find the boundary of mammogram region in DBA case, but sensitive to high gray values of background noise.

2.2 Removing specific background noise

The need of this step is due to the background noise especially in the Lumisys case, where the gray values of background are much than zero and consist of impulse noise and uneven illumination. This step is sensitive to the different database, may need adjustment according to the given samples. There are three operations during this step,

removing impulse noise along the vertical direction,

removing uneven illumination along horizontal direction,

cutting mammogram region roughly from the image.

The test results shows that our preprocessor runs successfully on images of both Lumisys database and DBA database if this step works well, which means our thresholding method is correct.

2.3 Generating and developing of the contour

A mask is used to served as the working area of each element of our developing contour, on which a moving direction is chosen to force the element to point to next position based on the lower-threshold and the gray values over this mask and those on each neighborhood mask with keeping the brighter area is always on the left hand side.

2.4 Adjustment of the contour.

Since we do not use upper-threshold during forming the contour, the contour may include the bright edge of the image of the original X-ray picture and something else. Here we point an advantage of our contour over other preprocessor, the latter of which have to find individual bright edges, lines and have a huge of task to remove them. Our contour has excluded all such signals outside the mammogram region, and has kept a few of remaining such signals inside the contour that is much easier to be located. In this step we can remove the bright wasting line according to the shape, smoothness of our contour.

2.5 Output of the results

The output are only the signals in the mammogram and all of the extrinsic signals outside of the mammogram were cut. One of significant advantages of our contour is that we can cut wasting area easily without any restriction on any directions, e.g. Some preprocessor have to do different removing in different directions.

3. The existing problem

Since we can not list all possible extrinsic signals, we always can encounter unlucky example. This usually happened on the change of database.

III. Basic algorithm design

Basic algorithm design includes nonlinear bank filter design, implementation and segmentation algorithm design and testing. *The nonlinear filter design is based on multiresolution techniques* which explained as follows.

Rationale for multiresolution techniques. These techniques provide a unique way of exploring the several levels of spatial redundancy existing on images. The term resolution is application dependent. For wavelet transform methods, frequency resolution properties of the mutiresolution wavelet transform are used. Decomposition is then a band-bass decomposition of the original image, and can be achieved by using a set of low pass(*linear)* filter bank with decreasing cut-off frequency. This decomposition, however, is both object size and shape dependent, where specific objects are not isolated in a subset of decomposed subimages; particularly at higher spatial resolution of less than 100 microns. These *linear* filter banks have to be replaced by multiresolution *nonlinear* filter banks which is less dependent on object size/shape or image gray scale characteristics. Nonlinear filters, such as order statistic (OS) are proposed for image enhancement based on our previous experience with OS filters. Multiresolution in this project is defined as size resolution. No prior information for the size of MCC suspicious area(s) is known. The size of this area(s) can be defined by comparison with a more or less isotropic window (a square in our case). This scheme helps to explore the decomposition of several levels of spatial target-objects to be detected. The multiresolution nonlinear filter bank for multiresolution decomposition with shape information can be achieved by our design. The resolution decrease is achieved by increasing the size of a square window mask. In our decomposition scheme, the filter F_m corresponding to the m^{th} level involves a square mask of (2m+1)x(2m+1). Thus, the five first levels are computed with squares of size 3x3, 5x5, 9x9, 17x17 and 33x33, respectively. Ideally, the filters should remove the objects which are smaller than a certain size and leave others unchanged. However, with a simple OS filter, it is not possible to achieve this. None of these filters allows perfect preservation of shapes, even if some are better than others for some purpose. To solve the problem of shape preservation, a stage called reconstruction is added after filtering. Its goal is to restore the original shape of the objects which have not been completely removed by the filtering process. The reconstruction stage proposed here is based on modifications and combinations of geodesic dilation and erosion.

Segmentation: We have tested 20 different segmentation algorithms on MCCs segmentation, such as Relaxation, Digital Desk - adaptive, Fuzzy sets, Otsu's method for grey level histograms, iterative selection, Johannsen Kapur method for using entropy, two histogram peaks, Minimum error and mean, Black percentage, Pun method for using entropy, two histogram peaks, etc. Recently, a novel segmentation algorithm was created based on the tests of the 20 different segmentation algorithms, which we named as "Iterative selection of Entropy of Fuzzy sete for Thresholdings (IEFT)". IEFT showed the best performance for the segmentation results for a small set of database. This method is modified from our existing method called Adaptive iterative thresholding and Entropy of Fuzzy set method. A paper is in progress on this issue.

IV. Conclusion

Hundreds of researchers in the past decade created hundreds of CAD methods for abnormal detection and classification in digitized/digital mammography, but the successes are very limited. The problems are summarized as follows; (1). Case dependent and date base dependent, (2). Digitizer dependent and resolution dependent, and (3). Manual justification of CAD parameters based on given data base. We are working on the solutions for the above problems.

V. References

1. Qian W " A Novel Hybrid Filter Architecture for Image Enhancement in Medical Imaging " Chapter in the Handbook of Medical Image Processing, 2000 by Academic Press.

2. Qian W, Xuejun Sun, Hong Liu and Robert Clark, "Wavelet-based image processing for digital mammography" Invited paper and Invited symposium speaker: Wavelet Application in Signal and Image Processing VIII, SPIE's International Symposium on Optical Science and Technology, July 30th -August 4th 2000, San Diego, USA

3. Qian W, Eshan Sheybani, Ravi Sankar, Dansheng Song, Xuejun Sun, Lin Zhang Hong Liu and Robert Clark, "High Speed Network for Telemammography", International Workshop in Digital Mammography, June 11-14, 2000, in Toronto, Canada

4. Qian W, Hong Liu, Lihua Li Robert Clark,"A Novel CAD Method for Mass Detection in Digital Mammography" IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, JULY 30 -- AUGUST 2, 2000, NEW YORK CITY.

Key research accomplishments

(1). Data base collection and truth file establishment for three sensors
(2). Well done the breast area segmentation for the mammograms
(3). Well done the basic algorithm design including nonlinear bank filter design, implementation and segmentation algorithm design and testing.

Reportable Outcomes

---Manuscripts, abstracts, presentations;

1. Qian W " A Novel Hybrid Filter Architecture for Image Enhancement in Medical Imaging " Chapter in the Handbook of Medical Image Processing, 2000 by Academic Press.

2. Qian W, Xuejun Sun, Hong Liu and Robert Clark, "Wavelet-based image processing for

digital mammography" Invited paper and Invited symposium speaker: Wavelet Application in Signal and Image Processing VIII, SPIE's International Symposium on Optical Science and Technology, July 30th -August 4th 2000, San Diego, USA

3. Qian W, Eshan Sheybani, Ravi Sankar, Dansheng Song, Xuejun Sun, Lin Zhang Hong Liu and Robert Clark, "High Speed Network for Telemammography", International Workshop in Digital Mammography, June 11-14, 2000, in Toronto, Canada

4. Qian W, Hong Liu, Lihua Li Robert Clark,"A Novel CAD Method for Mass Detection in Digital Mammography" IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, JULY 30 -- AUGUST 2, 2000, NEW YORK CITY.

---degrees obtained that are suported by this award;

Lin Zhang, working on Master degree, will be finished in December of 2000

----Funding applied for based on work supported by this award;

Conclusion

Hundreds of researchers in the past decade created hundreds of CAD methods for abnormal detection and classification in digitized/digital mammography, but the successes are very limited. The problems are summarized as follows; (1). Case dependent and date base dependent, (2). Digitizer dependent and resolution dependent, and (3). Manual justification of CAD parameters based on given data base. We are working on the solutions for the above problems.

References

1. Qian W " A Novel Hybrid Filter Architecture for Image Enhancement in Medical Imaging " Chapter in the Handbook of Medical Image Processing, 2000 by Academic Press.

2. Qian W, Xuejun Sun, Hong Liu and Robert Clark, "Wavelet-based image processing for digital mammography" Invited paper and Invited symposium speaker: Wavelet Application in Signal and Image Processing VIII, SPIE's International Symposium on Optical Science and Technology, July 30th -August 4th 2000, San Diego, USA

3. Qian W, Eshan Sheybani, Ravi Sankar, Dansheng Song, Xuejun Sun, Lin Zhang Hong Liu and Robert Clark, "High Speed Network for Telemammography", International Workshop in Digital Mammography, June 11-14, 2000, in Toronto, Canada

4. Qian W, Hong Liu, Lihua Li Robert Clark,"A Novel CAD Method for Mass Detection in Digital Mammography" IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, JULY 30 -- AUGUST 2, 2000, NEW YORK CITY.

Appenices

1. Key research accomplishments

(1). Data base collection and truth file establishment for three sensors
(2). Well done the breast area segmentation for the mammograms
(3). Well done the basic algorithm design including nonlinear bank filter design, implementation and segmentation algorithm design and testing.

2. Reportable Outcomes

---Manuscripts, abstracts, presentations;

1. Qian W " A Novel Hybrid Filter Architecture for Image Enhancement in Medical Imaging " Chapter in the Handbook of Medical Image Processing, 2000 by Academic Press.

2. Qian W, Xuejun Sun, Hong Liu and Robert Clark, "Wavelet-based image processing for digital mammography" Invited paper and Invited symposium speaker: Wavelet Application in Signal and Image Processing VIII, SPIE's International Symposium on Optical Science and Technology, July 30th -August 4th 2000, San Diego, USA

3. Qian W, Eshan Sheybani, Ravi Sankar, Dansheng Song, Xuejun Sun, Lin Zhang Hong Liu and Robert Clark, "High Speed Network for Telemammography", International Workshop in Digital Mammography, June 11-14, 2000, in Toronto, Canada

4. Qian W, Hong Liu, Lihua Li Robert Clark,"A Novel CAD Method for Mass Detection in Digital Mammography" IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, JULY 30 -- AUGUST 2, 2000, NEW YORK CITY.

---degrees obtained that are suported by this award;

Lin Zhang, working on Master degree, will be finished in December of 2000

----Funding applied for based on work supported by this award;
Proposal Submitted to Agency: NIH, R21 on 10/01/99
A Novel CAD Method for Telemammography

Proposal Submitted to Agency: NIH, R01 on 02/01/00
E-Mammography, A high speed network for Telemammography